Size and composition of the lexicon in prematurely born very-low-birth-weight and full-term Finnish children at two years of age

Stolt, Suvi

2007

Stolt , S , Klippi , A , Launonen , K , Munck , P , Lehtonen , L , Lapinleimu , H , Haataja , L & the PIPARI Study Group 2007 , ' Size and composition of the lexicon in prematurely born very-low-birth-weight and full-term Finnish children at two years of age ' , Journal of Child Language , vol. 34 , no. 2 , pp. 283-310 . https://doi.org/10.1017/S0305000906007902

http://hdl.handle.net/10138/309068
https://doi.org/10.1017/S0305000906007902

publishedVersion

J. Child Lang. **34** (2007), 283–310. © 2007 Cambridge University Press
doi:10.1017/S0305000906007902 Printed in the United Kingdom

Size and composition of the lexicon in prematurely born very-low-birth-weight and full-term Finnish children at two years of age*

SUVI STOLT
University of Helsinki and Turku University Central Hospital

ANU KLIPPI AND KAISA LAUNONEN
University of Helsinki

AND

PETRIINA MUNCK, LIISA LEHTONEN, HELENA LAPINLEIMU, LEENA HAATAJA AND THE PIPARI STUDY GROUP**
Turku University Central Hospital

(*Received* 7 *October* 2005. *Revised* 27 *October* 2006)

ABSTRACT

This paper focuses on the aspects of the lexicon in 66 prematurely born very-low-birth-weight and 87 full-term Finnish children at 2;0, studied using the Finnish version of the *MacArthur Communicative Developmental Inventory*. The groups did not differ in vocabulary size. Furthermore, the female advantage in vocabulary size was not seen in preterm children. The overall shapes of the trajectories for the main lexical categories as a function of vocabulary size were highly similar in both groups and followed those described in the literature. However, there were significant differences in the percentage of nouns and

[*] This study was supported by grants from the Emil Aaltonen Foundation, the Friends of Paediatric Clinics in University Hospitals in Finland Fund and the Neonatal Research in Southwestern Finland Fund.

[**] The Pipari Study Group: Satu Ekblad, Eeva Ekholm, Pentti Kero, Jarkko Kirjavainen, Riikka Korja, Harry Kujari, Hanna Manninen, Jaakko Matomäki, Jonna Maunu, Pekka Niemi, Ville Paasonen, Pertti Palo, Riitta Parkkkola, Jorma Piha, Annika Pihlgren, Liisi Rautava, Päivi Rautava, Milla Reiman, Hellevi Rikalainen, Katriina Saarinen, Elina Savonlahti, Matti Sillanpää, Päivi Tuomikoski-Koiranen, Tuula Äärimaa, and the following authors of the present article: Leena Haataja, Helena Lapinleimu, Liisa Lehtonen, Petriina Munck and Suvi Stolt. Address for correspondence: Suvi Stolt, Niittykatu 11, 20720 Turku, Finland. tel: +358-50-543 9021; e-mail: suvi.stolt@helsinki.fi

grammatical function words between the two groups. The results suggest that prematurity 'cuts off' the female advantage in vocabulary development. Furthermore, it also seems that there are differences between prematurely born and full-term children in the composition of the lexicon at 2;0. The findings support the universal sequence in the development of lexical categories.

Prematurely born (born before 37 weeks of gestation) very-low-birth-weight (VLBW, birth weight <1501 g) children are at risk of language deficits. For example, Mikkola *et al.* (2005) report in a study of all live born extremely-low-birth-weight (ELBW, birth weight <1000 g) children (live born: $n=351$, survived until the age of 5 years: $n=206$, number of assessed children: $n=172$) born in Finland during the two-year period 1996–1997, that language measures in developmental neuropsychological assessment (NEPSY test; Korkman, Kirk & Kemp, 1997) at 5;0 were significantly poorer in ELBW children when compared with normal population means. Similar findings have also been reported in other studies (e.g. Luoma, Herrgård, Martikainen & Ahonen, 1998; Wolke & Meyer, 1999). However, only a few studies focusing on aspects of early language acquisition in VLBW children have been done in the last ten to fifteen years (see however Casiro, Moddemann, Stanwick, Panikkar-Thiessen, Cowan & Cheang, 1990; Menyuk, Liebergott & Schultz, 1995; Riitesuo, 2000; Rvachew, Creighton, Feldman & Sauve, 2005). Detailed information on early language skills in this population would be important in order to recognize those in need of support in their language development as early as possible.

The development of the first lexicon in VLBW children has been investigated in only a few studies. Menuyk *et al.* (1995) followed the acquisition of vocabulary (comprehension of the first 10, 50 and 100 words, production of 10 and 50 words) of 26 premature (birth weight between 794 and 2500 g), including 12 VLBW, children, and 27 full-term children with the help of parental diaries. They found no difference between the groups of prematurely born and full-term children in early lexical acquisition. However, when the vocabulary development of 12 VLBW children was studied separately, VLBW children acquired their first 10 words significantly later than full-term children. Jansson-Verkasalo (2003) studied the language skills of 17 VLBW and 17 matched controls at 2;0 and 4;0. Among other things, the size of the lexicon was measured at 2;0 using the Finnish version of the *MacArthur Communicative Developmental Inventory* (CDI; Fenson, Dale, Reznick, Bates, Thal & Pethick, 1994; Lyytinen, 1999). The vocabulary size in VLBW children was smaller than that of full-term children, but the difference was not significant. Rvaschew *et al.* (2005) found that VLBW children ($n=13$) with chronic lung disorder,

a diagnosed bronchopulmonary dysplasia, had significantly smaller expressive vocabulary sizes at 18 months than healthier preterm ($n=9$) and full-term infants ($n=10$). Furthermore, in the study of Sansavini, Guarini, Alessandroni, Faldella, Giovanelli & Salvioli (2006), the lexical and grammatical development was assessed in the group of 77 low-birth-weight (<1600 g) and 22 full-term children at 2;6. The groups did not differ in size or composition of lexicon. However, the composition was not studied in detail, and the control group in relation to the study group was small. To conclude, the studies concentrating on the early lexicon in VLBW children done so far are few, and there is a need for further studies.

The predictive value of the early lexicon

The question of the development of the first lexicon in VLBW children is important, as it has been shown that the early lexicon may have predictive value for later language skills (e.g. Oliver, Dale & Plomin, 2004). Children who have a small lexicon size in the beginning of the third year of life, but are otherwise developing normally, have been classified as children with slow expressive language development (Paul, 2001), or as late talkers (e.g. Weismer, 2001). In follow-up studies (Oliver *et al.*, 2004; Paul, 2001; Weismer, 2001), it has been found that approximately half of these children continue to have language problems throughout their preschool period. The deficit found in vocabulary development in the beginning of the third year of life may appear later in the areas of phonology, morphology, syntax or narrative skills (Paul, 2001; Rescorla, 2005; Weismer, 2001). The predictive value of the late talking history may also be different in at-risk groups than in children with no risk for language problems. Lyytinen *et al.* (2001) found that late talkers in a group of at-risk for dyslexia were still delayed on language comprehension and production at 3;5, while late talkers in the control group performed at age-level expectations.

The composition of the lexicon may also differ in at-risk groups from that of normally developing children with no risk for language problems. Koster, Been, Krikhaar, Zwarts, Diepstra & Van Leeuwen (2005) reported that Dutch children with a familial risk of dyslexia had significantly fewer closed class words and verbs in their lexicon than controls at 1;7 when the lexicon size was controlled. However, it is not known yet whether the differences in the composition of the lexicon have predictive value for later language skills.

The development of the size and composition of the lexicon

Normally developing full-term children have acquired a basis for a lexicon of their first language by 2;0. If measured with a structured parental rating

method, for example the CDI, the median value for the lexicon size is between 200 and 400 words in the vocabularies of children growing up in different linguistic environments (Bates *et al.*, 1994; Lyytinen, 1999; Maital, Dromi, Sagi & Bornstein, 2000). The variation in the size of the vocabulary at this age is wide (Bates *et al.*, 1994; Lyytinen, 1999). However, the majority of the children have passed the period of the first 50 words at the age of 2;0. Bates *et al.* (1994) report that 10% out of 1803 English-speaking children produced fewer than 57 words at 2;0 (compare Menyuk *et al.*, 1995).

In the beginning of language acquisition there is variation not only in the lexicon size, but also in the composition of vocabulary. In addition to the stylistic variation (i.e. referential and expressive children) first described by Nelson (1973), there has been a discussion in the literature related to the composition of the early lexicon (Bates *et al.*, 1994; Bornstein & Cote, 2004; Caselli *et al.*, 1995; Caselli, Casadio & Bates, 1999; Jackson-Maldonado, Thal, Marchman, Bates & Gutierrez-Clellen, 1993; Kauschke & Hofmeister, 2002; Maital *et al.*, 2000). According to the universal proposal set by Gentner (1982), nouns are more prevalent than other word classes, especially verbs, in the early stages of language development. Based on the studies on English and Italian children, Gentner's original noun–verb proposal has been expanded to a four-stage model of lexical development. The model argues for a universal sequence from routine words to nouns, then to verbs, and after that, to grammar (Bates *et al.*, 1994; Caselli *et al.*, 1995; Caselli *et al.*, 1999). In the very beginning of lexical development (from 0 to 10 words), children use mainly verbal routines or vocal conventions in familiar situations (i.e. words such as sound effects for vehicles or animals, or names for people with different meanings). When the lexicon size is about 50 to 200 words, the number and proportional share of nouns clearly increases, but as soon as children acquire more words in their lexicon, this proportional share starts to decrease. In small vocabularies, the predicates (i.e. verbs and adjectives) are very rare. The number of these words starts to develop only after the lexicon size exceeds 100 words. The number of grammatical function words (i.e. closed class words) increases only when children have acquired a lexicon size of between 300 and 500 words. The developmental changes in the composition of the early lexicon are based on universal constraints of perception, memory, production and the availability of cognitive and conceptual structures that underlie human languages (Caselli *et al.*, 1999). They reflect a shift from reference to predication and to grammar (Bates *et al.*, 1994; Caselli *et al.*, 1995; Caselli *et al.*, 1999).

The overall shapes of the trajectories for different lexical categories as a function of vocabulary size have been found to be very similar, at least in the lexicons of Spanish (Jackson-Maldonado *et al.*, 1993), English (Bates

et al., 1994), Italian (Caselli *et al.*, 1999) and Hebrew (Maital *et al.*, 2000) children (for a comparison see Kauschke & Hofmeister, 2002). In addition to these similarities, some differences have been reported. In a comparison study of American English- and Italian-speaking children (Caselli *et al.*, 1999), it was found that the percentage of social terms (i.e. onomatopoeic words, names for people, games and routines) was higher, and the percentage of nouns was smaller, at the beginning of vocabulary development (< 50 words) in the lexicons of Italian than in American children. Moreover, the trajectory for closed class words gradually increased during the early lexical development (from < 50 words to > 600 words) of Italian children, while there was no proportional increase for these words in the lexicons of American children before the lexicon size was more than 400 words. Caselli *et al.* (1999) suggest that these findings reflect both cultural differences between the countries and morphological differences in the target language. The studies on the composition of the lexicon in Finnish children have been limited to small qualitative studies thus far (Nieminen, 1991), suggesting only very roughly that there are nouns and verbs as well as onomatopoeic words in the lexicon of Finnish children during the second year of life.

Factors associated with the language and lexical development

The group of VLBW children is a heterogeneous one. Because of their high risk at birth, biological factors have more impact on their cognitive development than in the full-term population. Some most typically reported factors associated with the language development of these children are overall cognitive development, whether the birth weight of a child is small for gestational age (SGA), gender and maternal education or socioeconomic status (SES) of the family. As far as we know, the impact of these biological and environmental factors on the early lexical development (i.e. before or at 2;0) in VLBW children has not been studied directly. However, some information can be gained from the studies concentrating on the overall language skills, areas of language skills other than the lexicon or on the language development of older children. Some studies report that language problems in VLBW children are mainly related to a low overall intelligence quotient (IQ; e.g. Mikkola *et al.*, 2005). This is not found in all studies, however. For example, in the study of Landry, Smith & Swank (2002), it was found that language difficulties in VLBW children appeared to be independent of their general cognitive problems. Further, the language outcome of SGA children has been reported to be significantly lower when compared to those prematurely born children with appropriate birth weight (e.g. Mikkola *et al.*, 2005), although this negative effect on the language skills of SGA has not been found in all studies (e.g. Casiro *et al.*, 1990).

Prematurely born boys have been reported to score lower than girls in language measures. Menyuk *et al.* (1995) found that boys scored lower in early two-word comprehension, and Luoma *et al.* (1998) that boys had a greater discrepancy between their performance and verbal IQ scores at 5;0. Furthermore, in the study of Sansavini *et al.* (2006), prematurely born boys had a significantly smaller lexicon size than girls at 2;6. However, this gender difference is not found in all studies. Boys born before 32 gestational weeks scored higher than girls in imitation of articulation patterns, imitation of sentences of differing grammatical complexity and word repetition tasks at 6;6 in a study by Jennische & Sedin (1999). Moreover, maternal education has been quite consistently reported to correlate positively to the language outcome of prematurely born children (e.g. Klebanov, Brooks-Gunn & McCormick, 1994), as has also been the SES of the family (e.g. Landry *et al.*, 2002). Sansavini *et al.* (2006) reported, however, that parental education had no effect to the lexicon size of prematurely born children (<1601 g) at 2;6. To conclude, both biological and environmental features have an effect on the language development of prematurely born children. It is unclear, however, what kind of influence these matters have on the very early lexical development of VLBW children.

In normally developing, full-term children the most typically reported factors associated with the variability in lexical development are gender and parents' education or social class. The female advantage in early lexical acquisition in full-term, normally developing children has been revealed in many studies (e.g. Fenson *et al.*, 1994; Bornstein, Leach & Haynes, 2004). Maternal education has been shown to have a positive effect on the early lexicon in some studies (e.g. Dollaghan *et al.*, 1999). This finding is inconsistent, however (e.g. Pan, Rowe, Spier & Tamis-LeMonda, 2004).

Relevant aspects of Finnish

Finnish is an agglutinative language in which grammatical and case relations are expressed primarily by the means of suffixes (Toivainen, 1997). Nominals (i.e. nouns, adjectives, pronouns and numerals) and participial forms of verbs are inflected with the help of fifteen cases. Infinitival forms of verbs can also be inflected for case, though this is more restricted. Verbal morphology for finite verbs can express voice (active, passive), person (first, second or third), number (singular, plural), tense (present, past) and mood (indicative, imperative, conditional, potential). In addition to the extensive nominal and verbal inflectional system, there are many morphophonological alternations in Finnish. For example, in the consonant gradation, a strong-grade form of the consonant is weakened in certain environments (e.g. *nu**kk**u-u* 'sleep-PRES+3SG', 'sleeps': *nu**k**un*

'sleep-PRES+1SG', 'I sleep'; Toivainen, 1997). The case marking and person inflection play a major role in the coding of syntactic roles. Objects and subjects are often distinguished from each other with the help of case marking, or verbal inflection is used to mark the subject (Helasvuo, to appear; Toivainen, 1997).

The rich morphological system may affect the lexical acquisition of Finnish children. Because of the intensive use of suffixes, Finnish words are relatively long. Furthermore, in the lexical acquisition process, Finnish children need to distinguish between base forms and inflections. This sets high demands for the auditory processing system of linguistic units. On the other hand, the different morphological marking of nouns and verbs may also help children to separate the different word classes. The use of intensive morphological inflections may be especially challenging for those who have difficulties with speech perception (Lyytinen & Lyytinen, 2004). Difficulties in the auditory processing systems have been found in VLBW children (Jansson-Verkasalo, 2003).

Aims of the present study

The primary aim of the present study is to gather detailed information on aspects of the lexicon in VLBW children at 2;0. The purpose is to ascertain whether the size of the lexicon in VLBW children differs from that of full-term healthy children, whether there are more children with very small vocabularies (< 50 words) in VLBW than in full-term children and whether the composition of the lexicon in these two groups differs from each other. Moreover, the aim is to gain information on the composition of the lexicon in full-term healthy Finnish children, as it is not specifically known how and in what order Finnish children learn different lexical categories in their early vocabulary acquisition. In addition, the effect of background variables on the lexicon size is studied in both groups.

METHOD

Participants

The size and composition of the lexicon were analyzed in a group of VLBW and healthy full-term children. All children were participants of a multidisciplinary follow-up study (the PIPARI study). The inclusion criteria of the PIPARI study for prematurely born children are: the birth weight of the child <1501 g, the parent understands Finnish or Swedish well enough to be able to complete the follow-up forms and the families live in the Turku University Hospital catchment area. The PIPARI study has been approved by the Ethical Committee of the Hospital District of Southwest Finland in December 2000 (VLBW children) and in September

2001 (full-term children). The VLBW children in the present study were born between January 2001 and December 2002. All prematurely born VLBW children meeting the inclusion criteria were invited, and all the families participated. From the total sample of 100 prematurely born infants, 17 (17%) died and four families (4%) dropped out of the study during the two-year follow-up. The full-term children were born at Turku University Hospital between November 2001 and March 2003. The first healthy (i.e. normal birth weight, no admissions in the neonatal intensive care unit) full-term (born >37 gestational weeks) boy and girl of the week were invited to join the study. In all, 117 families agreed to participate. During the two-year follow-up period two (1·7%) families dropped out.

Data collection

At 2;0, the children's cognitive development was measured using the *Bayley Scales of Infant Development* (BSID II; Bayley, 1993). The corrected age was used for the VLBW children. The age correction was done by counting the child's age from the expected date of delivery. The date was determined from ultrasound examination done in the early stage of the pregnancy, or from the maternal dates of the last menstrual period. The use of corrected age during the first years of life in VLWB children is a commonly accepted method (e.g. Wolke & Meyer, 1999), and is used to compare the development of VLBW and full-term children at the same developmental age. At the two-year appointment, the families received the Finnish version of the CDI (words and sentences), and they were asked to complete and return it within two weeks by post. In the VLBW group, 73 (92%) inventories were returned. From those inventories, 7 were left out of the present analysis because the families reported the use of a language or languages other than Finnish at home. In the group of full-term children, 99 (86%) inventories were returned. From those 11 were left out of the present analysis because the language used in the family was not Finnish, and 1 because it was not properly completed. The final sample consisted of 66 inventories of VLBW and 87 inventories of healthy full-term children growing up in Finnish-speaking monolingual families. The mean corrected age of the VLBW children was 2;0·11 (*S.D.* 17 days; chronological age: 2;3·1, *S.D.* 23 days), and the mean age of the full-term children 2;0·13 (*S.D.* 12 days) at the time the CDI form was completed.

The hearing threshold of 29 VLBW children was measured with brainstem auditory evoked potential between 0;0 and 0;2 corrected age. One child was sent for further examination and was found to have a hearing impairment (see below). The hearing threshold of the other 28 children was found to be normal for this age. Furthermore, the otoacustic emission was

used to measure the hearing of two VLBW children. In addition, and according to normal clinical routines, the VLBW children were followed intensively during the first two years, and sent for audiologist examination if hearing impairment was suspected. The medical records of all VLBW children were checked at 2;0 to see whether hearing impairment had been diagnosed. According to the medical records, none, except the one already mentioned, had a diagnosed hearing impairment. The hearing of full-term children was not measured in a formal way. However, none was using a hearing aid or had a diagnosed hearing impairment at the 2;0 appointment.

The VLBW and full-term children were divided into two groups according to the mental developmental index (MDI) in the BSID II. Group 1 included children with an MDI $\geqslant$ 85, 61 VLBW and 84 full-term children. In Group 2 there were 5 VLBW and 3 full-term children with an MDI $<$ 85 (-1 *S.D.*). Thus, we compared the lexicon of those VLBW and full-term children developing cognitively according to their age (Group 1), and the lexicon of those children not having achieved the age-appropriate cognitive skills at 2;0 (Group 2). The group division was done to get as clinically relevant information of the lexicon of the VLBW children as possible. The group of VLBW children is a heterogeneous one, and the cognitive impairments in these children are common (e.g. Wolke & Meyer, 1999). In order to get precise information of the lexicon size and composition of those children developing cognitively according to their age, as well as those who are not, the cognitive development was controlled (compare e.g. Luoma *et al.*, 1998; Wolke & Meyer, 1999).

Group 1. Although children in Group 1 were developing cognitively according to their age, the MDI values of the prematurely born VLBW and full-term children differed ($Z = -2{\cdot}05$, $p = 0{\cdot}04$). The basic education of the mothers was classified into three categories (1 = schooling interrupted before nine years at comprehensive school, 2 = nine years at comprehensive school finished, 3 = nine years at comprehensive school and three senior grades at secondary school). The maternal education data of 3 children in the VLBW group and 4 in the full-term group were missing. No significant difference was seen between the two groups in the basic education of the mothers ($\chi^2 = 0{\cdot}65$, df = 1, $p = 0{\cdot}42$) or in gender ($\chi^2 = 2{\cdot}88$, df = 1, $p = 0{\cdot}09$). The information of the mothers' basic education and the gender of the children is shown in Table 1.

Group 2. In Group 2, there were 5 VLBW and 3 full-term children. The MDI values in the 5 VLBW children varied between 50 and 80, and in the group of full-term children between 72 and 84 (see Table 1). Two VLBW children had cerebral palsy (CP), and one child with CP also had a symmetrical hearing impairment (hearing thresholds without hearing aids at the level 55–75 dB, and at the level of 30 dB with hearing aids).

TABLE 1. *Child and parent characteristics of prematurely born very-low-birth-weight (VLBW) and full-term (FT) children. In Group 1 there are children with mental developmental index ⩾85, and in Group 2 children with mental developmental index <85 on the Bayley Scales of Infant Development. Mean values (standard deviations) and minimum–maximum values are presented. If median values or percentages are used they are marked separately*

	VLBW children		***FT children***	
	Group 1 *n*=61		**Group 1 *n*=84**	
Birth weight (g)	1067 (263)	400–1500	3682 (460)	2790–4980
Gestational age in weeks	28 (3)	24–35	40 (1)	37–42
Apgar (median)	7	1–9	9	8–10
SGA children	22 (36%)/61		0/84	
MDI	107 (11)	86–128	111 (10)	88–128
Females/Males	24 (39%)/ 37 (61%)		45 (54%)/ 39 (46%)	
Mother's education				
Less than 9 years	0		0	
9–12 years	20 (33%)/61		33 (39%)/84	
over 12 years	38 (62%)/61		47 (56%)/84	
	Group 2 *n*=5		**Group 2 *n*=3**	
Birth weight (g)	872 (426)	525–1475	4178 (843)	3285–4960
Gestational age in weeks	27 (3)	23–31	39 (1)	39–40
Apgar (median)	4	2–7	9	9–9
SGA children	1 (20%)/5		0	
MDI	68 (13)	50–80	80 (7)	72–84
Females/Males	2 (67%)/3		0/3	
Mother's education				
Less than 9 years	0		0	
9–12 years	3 (60%)/5		2 (67%)/3	
over 12 years	2 (40%)/5		1 (33%)/3	

SGA = small for gestational age, weight < -2 *S.D.* from the mean of Finnish growth charts. The percentages of the mother's basic education do not total 100% in children of Group 1, because the data of 3 (5%) prematurely and 4 (5%) full-term children's mothers were missing.

Measures

To study the size and composition of the lexicon at 2;0, we used a structured parental rating method, the standardized Finnish version of the CDI (Lyytinen, 1999). In the normative study of the Finnish version of the CDI, the vocabulary development of 95 full-term, healthy children was followed at ages 1;0, 1;2, 1;6, 2;0 and 2;6. In this process all items on the

word lists were screened for linguistic and cultural relevance (Lyytinen, 1999). The criteria for a word are specified in the instructions given to parents together with the actual word lists in the Finnish version of the CDI. Only the words a child uses spontaneously (i.e. no imitated words) and repeatedly are accepted. The English version of CDI has been shown to be a reliable and valid method in many studies (e.g. Bornstein *et al.*, 2004; Pine, Lieven & Rowland, 1996). The Finnish version has been found to be reliable as well (Lyytinen, 1999). In the normative study of the Finnish version of the CDI, the concurrent correlation between the lexicon size in the CDI and the expressive scale in the Reynell Developmental Language Scale at 1;6 was high ($r = 0{\cdot}85$, $p \leqslant 0{\cdot}001$). At 2;0, the lexicon size correlated significantly and positively with the MDI value in the BSID ($r = 0{\cdot}70$, $p \leqslant 0{\cdot}001$) (Lyytinen, 1999). The CDI is also well suited to the context of the present study (Pine *et al.*, 1996), i.e. to the comparison of individual differences.

There are 595 items presented in 20 categories in the Finnish version of the CDI (words and sentences). Word categories on the list are sound effects and animal sounds (13 items), animals (38 items), vehicles (12 items), toys (15 items), food and drink (58 items), clothing (28 items), body parts (24 items), furniture and rooms (32 items), small household items (48 items), outside things and places to go (38 items), people (24 items), games and routines (22 items), action words (106 items), words about time (12 items), descriptive words (54 items), pronouns (24 items), question words (8 items), prepositions and locations (20 items), quantifiers (9 items) and connectives (10 items).

The words that the children use are defined according to the categories of adult language in the present study, as has been done in recent studies on vocabulary composition (e.g. Bates *et al.*, 1994; Bornstein & Cote, 2004; Kauschke & Hofmeister, 2002). It is known, however, that in the beginning of lexical acquisition children may use the same word with many different meanings, for example the word *hot* may function as an adjective or as a noun. One may argue what the classification into adult language categories tells about child language. It has been shown, however, that children acquire lexical categories at a different rate during the early stages of vocabulary development (Bates *et al.*, 1994; Bornstein & Cote, 2004; Caselli *et al.*, 1999; Kauschke & Hofmeister, 2002; Maital *et al.*, 2000). This can be seen to reflect their ability to handle different types of adult speech categories in the language input around them (Bates *et al.*, 1994; Caselli *et al.*, 1999). Thus, words categorized according to adult language can provide qualitative information about how children acquire different kinds of linguistic units around them.

Data handling. The values of Group 1 were analyzed as follows. The number of words produced by each child and reported by the mother was

counted, and the size of the vocabulary of the VLBW and full-term children was compared statistically. To compare the composition of lexicon in VLBW and full-term children, children were divided into 5 groups based on their vocabulary size (<50 words: 6 VLBW and 11 full-term children; 51–174 words: 14 VLBW and 16 full-term children; 175–300 words: 17 VLBW and 16 full-term children; 301–424 words: 8 VLBW and 25 full-term children; and >425 words: 16 VLBW and 16 full-term children). The grouping was based roughly on intervals similar to those in other vocabulary composition studies (Bates *et al.*, 1994; Caselli *et al.*, 1995, 1999; Koster *et al.*, 2005). There was no significant difference between VLBW and full-term children in the distribution over vocabulary size groups ($\chi^2 = 6{\cdot}92$, df = 4, $p = 0{\cdot}14$).

The composition analysis focused on four lexical categories: social terms, common nouns, predicates and grammatical function words. The category of social terms was the combined category of three categories of the CDI: sound effects and animal sounds, people and games and routines (59 items, 9·9% of the checklist). All noun categories on the CDI list serving a clear naming function were combined into one common noun category. The categories were animals, vehicles, toys, food and drink, clothing, body parts, furniture and rooms, and small household items (255 items, 42·9% of the checklist). Some potential nominals, especially the words included in the people or games and routines categories, were excluded from the category of nouns because previous studies have suggested that they may follow a different developmental course from 'true nominals' (Bates *et al.*, 1994; Caselli *et al.*, 1995). The predicate category included action words (i.e. verbs, 106 items, 17·8% of the checklist) and descriptive words (i.e. states and attributes, 54 items, 9% of the checklist). Grammatical function words (i.e. closed class words, 71 items, 11·9% of the checklist) included 5 categories on the CDI form: pronouns, question words, prepositions and locations, quantifiers and connectives. The categorization system used in the present study is very much the same as that used in other studies on the composition of the early lexicon (Bates *et al.*, 1994; Bornstein & Cote, 2004; Caselli *et al.*, 1995, 1999; Jackson-Maldonado *et al.*, 1993; Koster *et al.*, 2005; Maital *et al.*, 2000).

The composition analysis was accomplished in two different ways. Firstly, the percentages of social terms, common nouns, predicates and grammatical function words were calculated based on the total number of words marked on the CDI form (Bates *et al.*, 1994; Caselli *et al.*, 1999; Maital *et al.*, 2000). Secondly, word opportunity scores were used (Bates *et al.*, 1994; Bornstein & Cote, 2004). That is, the percentages of the four listed lexical categories were calculated based on the total number of items in that category on the CDI list (i.e. the options the mother was given when

completing the CDI form, e.g. what percentage of the 59 social terms words were reported).

Two different calculation styles were used for the following reasons. The aim of the present study was to get as comparable information on the composition of VLBW children and full-term Finnish children as possible with regard to the information already known. Calculating the percentages based on the total number of vocabulary items has been used in other studies of the composition of early vocabulary (Bates *et al.*, 1994; Caselli *et al.*, 1999; Maital *et al.*, 2000). However, there are differences in the absolute number of items within each category in the CDI (for the explanation, see Bates *et al.*, 1994; Caselli *et al.*, 1995). To make comparisons across word classes as equal as possible, we also used word opportunity scores (Bates *et al.*, 1994; Bornstein *et al.*, 2004; Pine *et al.*, 1996).

The percentages of social terms, common nouns, predicates and grammatical function words were calculated for each child in two different ways, as described earlier. The percentages of the listed lexical categories (mean values of the subgroup) in the lexicons of the VLBW and full-term children of similar sizes were compared statistically.

The values of Group 2 were analyzed separately. The composition analysis was done similarly to that of the children in Group 1. However, because of the small number of participants, the data was analyzed with descriptive numbers only.

Statistics. The data were first examined for normality using the Shapiro-Wilk's test. Comparisons between two normally distributed variables were done using an Independent Samples *t*-test. The Mann–Whitney *U* test was used for non-normal distributions. A Pearson's correlation coefficient was used to assess the association between vocabulary size and continuous background variables, and a chi-square test to analyze whether there were significant differences in categorical variables between the VLBW and full-term children. Statistical analyses were performed with SPSS for Windows (12.0, SPSS Inc.). Differences were considered statistically significant if the p-value was below 0.05. All significance tests were 2-tailed.

RESULTS

Vocabulary size. In Group 1 there was a high variability in productive vocabulary development in VLBW and in full-term children when examined as a function of the lexicon size. The range extended from 8 to 574 with a median of 244 words in the VLBW group (*M* 265, *S.D.* 170), and from 5 to 581 with a median of 281 words (*M* 272, *S.D.* 161) in the full-term children. The difference between the two groups was not significant ($Z = -0{\cdot}28$, $p = 0{\cdot}78$).

In Group 2, there was a clear difference in the median values between the VLBW and full-term children. The median value was 16 words in the group of 5 VLBW children (*M* 33, *S.D.* 45, min. 4, max. 111), and 72 words in the group of three full-term children (*M* 52, *S.D.* 36, min. 11, max. 73).

When the size of the lexicon of all children (Groups 1 and 2 together) was examined, the range extended from 4 to 574 with a median of 235 words in the VLBW group (*M* 247, *S.D.* 175), and from 5 to 581 with a median of 278 words (*M* 264, *S.D.* 163) in the full-term children. The difference between the groups was not significant ($Z = -0{\cdot}68$, $p = 0{\cdot}496$).

Six out of the 61 VLBW children (9·8%) and 11 out of the 84 full-term children (13·1%) had lexicons smaller than 50 words at 2;0 in Group 1. In Group 2, 4 VLBW children and 1 full-term child had vocabularies smaller than 50 words. When the values of all the children (Group 1 and 2 together) were assessed, the percentage of the children with a vocabulary <50 words was higher in the prematurely born children than in the full-term children (10 out of 66 VLBW children, 15·2%, and 12 out of 87 full-term children, 13·8%). This difference was not significant ($\chi^2 = 0{\cdot}06$, df = 1, $p = 0{\cdot}81$).

Vocabulary composition. The percentages, calculated on the total number of vocabulary items, of social terms, common nouns, predicates and grammatical function words in the different vocabulary size groups, were highly similar in the vocabularies of the VLBW and full-term children in Group 1 (see Table 2). Thus, the overall shapes of the trajectories for the lexical categories listed were very similar in the lexicons of the VLBW and full-term children (see Figure 1). The proportional share of social terms was high in small vocabularies (<50 words), but the percentages decreased as soon as the children had acquired their first 50–175 words. The clear increase of common nouns from 0 to 100 words changed to a proportional decrease after the children had achieved an approximate number of 100–150 words in their lexicons. The proportion of predicates increased steadily, taking the greatest gains between 51–174 and 175–300 words. The proportion of grammatical function words slowly increased all the way from the very small vocabularies up to a lexicon size of nearly 600 words.

Two significant differences between the VLBW and full-term children in Group 1 were found in the vocabularies >425 words. The percentage of common nouns was higher ($t(30) = 3{\cdot}56$, $p = 0{\cdot}001$), and the percentage of grammatical function words lower ($t(30) = -3{\cdot}15$, $p = 0{\cdot}004$) in the lexicons of the VLBW children than in those of the full-term children. In addition, it was possible to see differences, yet not statistically significant, between the two groups with very small lexicons (<50 words). There were differences in the percentages for social terms (mean values of the subgroups: 61% in the lexicon of the VLBW children, 53% in the full-term children), in the percentage of common nouns (25% in the VLBW children, 32% in the full-term children) and in the percentage of

TABLE 2. *Percentages of social terms, common nouns, predicates and grammatical function words calculated on the total number of vocabulary items at five vocabulary size categories in the lexicons of prematurely born very-low-birth-weight (VLBW) and full-term children (FT). Mean values, standard deviations (S.D.) and minimum–maximum values of the subgroups are presented. Comparisons between groups were done using the Independent Samples t-test or the Mann–Whitney U-test (U)*

	VLBW children n=61		***FT children n=84***		***Group comparison***
	Mean (S.D.)	Min.–Max.	Mean (S.D.)	Min.–Max.	*p*-value
	<50 words		*<50 words*		
Social terms	61 (17)	50–94	53 (19)	37–100	0·216 *u*
Nouns	25 (15)	0–39	32 (14)	0–49	0·387
Predicates	9 (4)	5–14	7 (6)	0–20	0·450
Grammatical function words	0 (0)	0–0	5 (7)	0–22	0·149 *u*
	51–174 words		*51–174 words*		
Social terms	23 (9)	14–45	21 (4)	15–28	0·822 *u*
Common nouns	56 (11)	33–67	56 (6)	47–70	0·914
Predicates	12 (7)	0–25	12 (4)	5–20	0·833
Grammatical function words	5 (2)	2–9	5 (2)	1–9	0·703
	175–300 words		*175–300 words*		
Social terms	13 (1)	10–15	13 (2)	10–18	0·983
Common nouns	55 (5)	45–64	52 (5)	45–62	0·129
Predicates	20 (4)	14–28	22 (3)	15–26	0·109
Grammatical function words	6 (2)	3–11	6 (2)	4–10	0·363 *u*
	310–424 words		*310–424 words*		
Social terms	11 (2)	9–14	11 (1)	9–13	0·634
Common nouns	50 (4)	46–59	50 (4)	42–58	0·553
Predicates	23 (4)	16–28	25 (4)	19–32	0·223
Grammatical function words	8 (2)	4–10	7 (2)	4–11	0·468
	>425 words		*>425 words*		
Social terms	10 (0)	10–11	10 (1)	9–13	0·956 *u*
Common nouns	47 (2)	44–50	45 (2)	41–48	0·001
Predicates	27 (2)	24–30	28 (2)	25–32	0·100
Grammatical function words	8 (1)	5–11	10 (1)	8–12	0·004

grammatical function words (none in the VLBW children, 5% in the full-term children).

Word opportunity scores for children in Group 1 showed a similar growth order of social terms, common nouns, predicates and grammatical function words to that found when percentages were calculated based on the

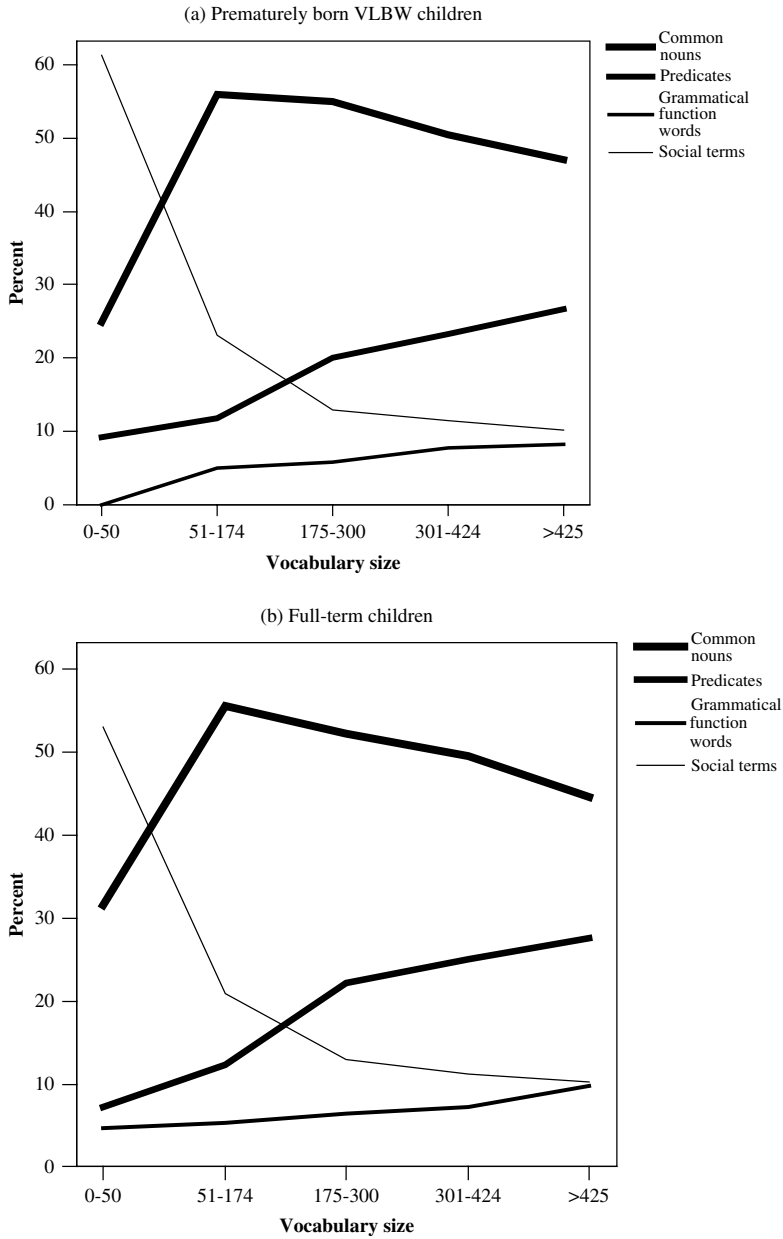

Fig. 1. Social terms, common nouns, predicates and grammatical function words as a proportion of total vocabulary size in the lexicons of prematurely born very-low-birth-weight (VLBW) children ($n=61$) and full-term children ($n=84$).

total number of vocabulary items (see Figure 2 and Table 3). Mothers of both groups completed the options for social terms first. They reported that their children began to learn common nouns at a faster rate than predicates or grammatical function words. Both groups of children had acquired approximately half of the common nouns listed on the CDI form when they had a vocabulary size of a 175–300 words. Approximately 50% of the predicates in the CDI had been marked after the vocabulary size was between 310 and 424 words. More than half of the grammatical function words on the Finnish CDI were acquired only after the vocabulary size had increased to >424 words.

One significant difference was found in the word opportunity scores between the VLBW and full-term children in Group 1. At >425 words, the VLBW children had fewer grammatical function words in their vocabulary than full-term children with a similar vocabulary size ($t(30) = -2{\cdot}30$, $p = 0{\cdot}029$).

Percentages for different lexical categories calculated based on the total number of vocabulary items for Group 2 are shown in Table 4, and word opportunity scores in Table 5. A roughly similar kind of growth order for different word categories was observed in the vocabularies of the children in Group 2 as found in the composition analysis of the children in Group 1.

Background characteristics. Associations between the lexicon size and the following background characteristics were analyzed: cognitive level, birth weight, whether the child was small for his or her gestational age (SGA, birth weight <−2 *S.D.* from the mean of Finnish growth charts) at birth, gender and the basic education of the mother.

A significant positive correlation was seen between vocabulary size and the cognitive level of the children (Group 1: $r = 0{\cdot}74$, $p < 0{\cdot}001$; all children: $r = 0{\cdot}73$, $p < 0{\cdot}001$). No correlation was found between birth weight and vocabulary size in Group 1 (VLBW children: $r = 0{\cdot}19$; full-term children: $r = -0{\cdot}05$). However, when all children were included, there was a significant positive correlation between birth weight and vocabulary size in VLBW children ($r = 0{\cdot}25$, $p = 0{\cdot}04$), but not in full-term children ($r = -0{\cdot}09$). There were 23 SGA children in the group of VLBW children (22 children in Group 1 and 1 child in Group 2). Growth retardation did not significantly influence the lexicon size at 2;0 (VLBW children in Group 1: $Z = -0{\cdot}45$, $p = 0{\cdot}66$; all VLBW children: $Z = -0{\cdot}49$, $p = 0{\cdot}62$).

The mean value of the lexicon size in the VLBW girls in Group1 was 262 words (*S.D.* 168, *Mdn* 236) and in the VLBW boys 266 (*S.D.* 173, *Mdn* 249). The difference between the VLBW girls and boys was not significant ($Z = -0{\cdot}01$, $p = 0{\cdot}99$). Correspondingly, the mean value for the full-term girls in Group 1 was 325 words (*S.D.* 152, *Mdn* 343) and for the boys 211 (*S.D.* 150, *Mdn* 175). This difference was significant ($Z = -3{\cdot}12$, $p = 0{\cdot}002$). When the values of all children (Group 1 and 2 together) were taken into

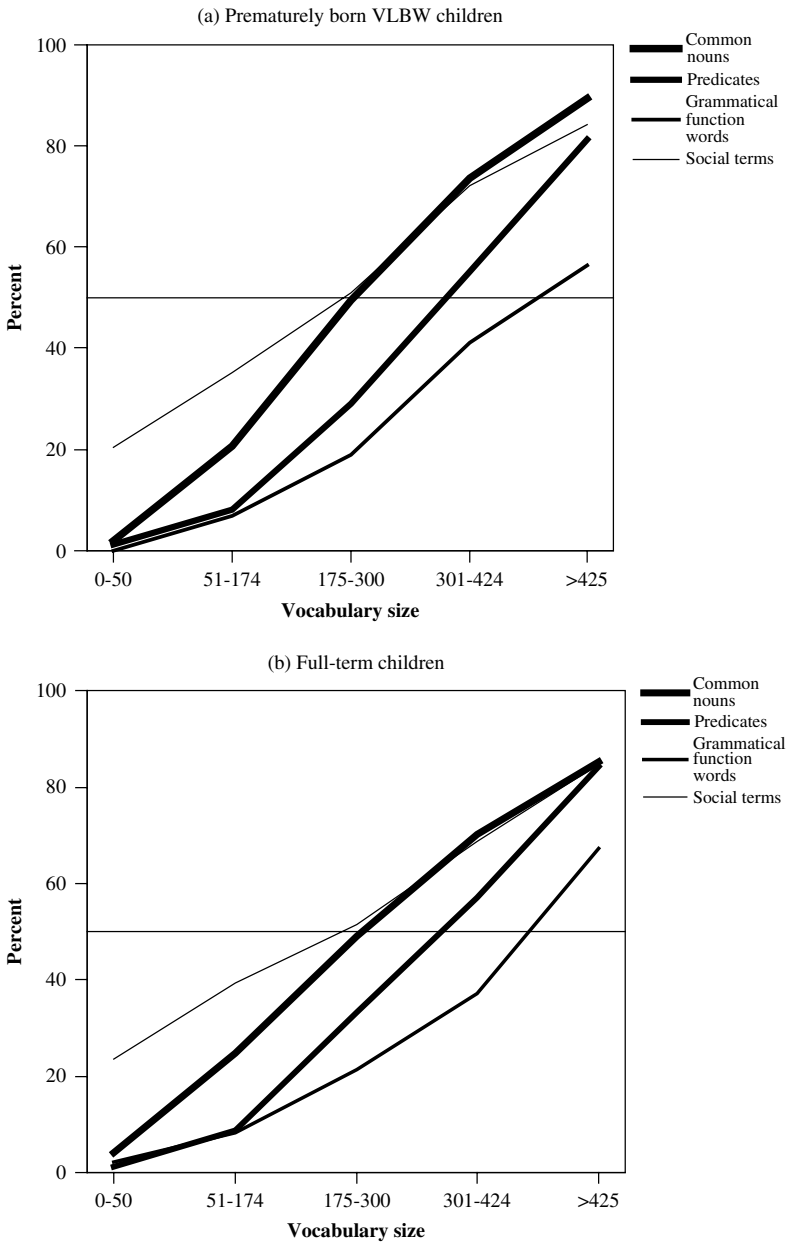

Fig. 2. Proportion of social terms, common nouns, predicates and grammatical function words on the checklist reported at each vocabulary size subgroup in the lexicons of prematurely born very-low-birth-weight children (VLBW, $n=61$) and full-term children ($n=84$).

TABLE 3. *Word opportunity scores for very-low-birth-weight (VLBW) and full-term (FT) children. Numbers presented are the proportions of words produced by the child relative to the total possible number of words in that category on a checklist*

	VLBW children n=61		***FT children n=84***		***Group comparison***
	Mean (*S.D.*)	Min.–Max.	Mean (*S.D.*)	Min.–Max.	*p*-value
	<50 words		*<50 words*		
Social terms	20 (7)	7–25	24 (8)	5–32	0·122 *u*
Common nouns	2 (2)	0–4	4 (3)	0–9	0·110
Predicates	1 (1)	1–3	1 (1)	0–3	0·961 *u*
Grammatical function words	0 (0)	0–0	2 (4)	0–11	0·149 *u*
	51–174 words		*51–174 words*		
Social terms	35 (13)	19–63	39 (11)	20–51	0·257 *u*
Common nouns	21 (7)	7–34	25 (8)	15–43	0·163
Predicates	8 (7)	0–26	9 (4)	3–17	0·240 *u*
Grammatical function words	7 (5)	1–18	8 (4)	1–14	0·294 *u*
	175–300 words		*175–300 words*		
Social terms	51 (9)	36–66	52 (7)	37–66	0·792
Common nouns	50 (8)	40–61	49 (9)	31–62	0·929 *u*
Predicates	29 (8)	15–44	33 (8)	19–44	0·147
Grammatical function words	19 (8)	11–45	21 (7)	13–34	0·191 *u*
	310–424 words		*310–424 words*		
Social terms	72 (9)	58–81	69 (7)	53–83	0·295
Common nouns	74 (7)	60–84	70 (8)	49–80	0·352 *u*
Predicates	55 (14)	31–73	57 (12)	39–80	0·680
Grammatical function words	41 (12)	21–56	37 (12)	18–66	0·431
	>425 words		*>425 words*		
Social terms	84 (7)	71–95	85 (9)	73–100	0·848
Common nouns	89 (6)	78–100	85 (8)	73–100	0·113
Predicates	81 (9)	64–93	84 (10)	68–100	0·373
Grammatical function words	56 (13)	35–90	67 (14)	45–90	0·029

Mean values, standard deviations (*S.D.*) and minimum–maximum values of the vocabulary size subgroups are presented. Comparisons between the subgroups were done using the Independent Samples t-test and the Mann–Whitney U-test (U).

consideration, the mean value for VLBW girls was 248 (*S.D.* 170, *Mdn* 215), and for the boys 247 words (*S.D.* 180, *Mdn* 243). The mean value for all full-term girls was 324 (*S.D.* 152, *Mdn* 343) and for the boys 200 (*S.D.* 150, *Mdn* 160). The difference between all VLBW girls and boys was not significant ($Z = -0{\cdot}203$, $p = 0{\cdot}839$), but the difference between all full-term girls and boys was ($Z = -3{\cdot}466$, $p = 0{\cdot}001$). Furthermore, the VLBW girls did not

TABLE 4. *Percentages of social terms, common nouns, predicates and grammatical–function words calculated on the total number of vocabulary items, for those prematurely born very-low-birth-weight (VLBW) and full-term (FT) children who were not developing cognitively according to their age. Median and minimum–maximum values of the vocabulary size subgroups are presented*

	VLBW children n=5		***FT children n=3***	
	Median	Min.–Max.	Median	Min.–Max.
	<50 words n=4		*<50 words n=1*	
Social terms	55	44–100	55	55–55
Common nouns	42	0–50	27	27–27
Predicates	0	0–6	18	18–18
Grammatical function words	0	0–0	0	0–0
	51–174 words n=1		*51–174 words n=2*	
Social terms	26	26–26	33	32–35
Common nouns	35	35–35	48	39–56
Predicates	17	17–17	7	3–11
Grammatical function words	14	14–14	8	4–11

differ from the full-term girls within Group 1 ($Z=-1{\cdot}58$, $p=0{\cdot}11$). Neither was the difference between the VLBW boys and the full-term boys significant (Group 1, $Z=-1{\cdot}34$, $p=0{\cdot}18$). However, when all children (Group 1 and Group 2) were included, the difference between the VLBW and full-term girls was nearly significant ($Z=-1{\cdot}93$, $p=0{\cdot}05$), but not the difference between the VLBW and full-term boys ($Z=-1{\cdot}02$, $p=0{\cdot}31$).

The basic education level of the mother was associated with the lexicon size in the group of VLBW children. The median value of the vocabulary size of VLBW children in Group 1 who had a mother with a basic education level of between nine and twelve years (*Mdn* 151, *M* 183, *S.D.* 155) was significantly lower than the median value of those children who had a mother with a basic education level of over twelve years (*Mdn* 298, *M* 307, *S.D.* 162, $Z=-2{\cdot}86$, $p=0{\cdot}004$). The trend was similar in the group of all VLBW children (9–12 years: *Mdn* 129, *M* 165, *S.D.* 153; over 12 years: *Mdn* 279, *M* 292, *S.D.* 170, $Z=-2{\cdot}97$, $p=0{\cdot}003$). A similar association was not found in the group of full-term children (Group 1, 9–12 years: *Mdn* 303, *M* 269, *S.D.* 157; over 12 years: *Mdn* 277, *M* 267, *S.D.* 160, $t(78)=0{\cdot}08$, $p=0{\cdot}94$; all full-term children, 9–12 years: *Mdn* 279, *M* 258, *S.D.* 160; over 12 years: *Mdn* 276, *M* 261, *S.D.* 162, $Z=-0{\cdot}04$, $p=0{\cdot}97$).

TABLE 5. *Word opportunity scores for those prematurely born very-low-birthweight (VLBW) and full-term (FT) children who were not developing cognitively according to their age. Numbers presented are the percentages of words produced by a child and calculated on the total possible number of words in that category on a checklist. Median and minimum–maximum values of the vocabulary size subgroups are presented*

	VLBW children n=5		***FT children n=3***	
	Median	Min.–Max.	Median	Min.–Max.
	<50 words n=4		*<50 words n=1*	
Social terms	9	3–32	10	10–10
Common nouns	2	0–5	1	1–1
Predicates	0	0–1	1	1–1
Grammatical function words	0	0–0	0	0–0
	51–174 words n=1		*51–174 words n=2*	
Social terms	49	49–49	41	39–42
Common nouns	15	15–15	14	11–16
Predicates	12	12–12	3	1–5
Grammatical function words	21	21–21	8	4–11

DISCUSSION

In this study, aspects of the lexicon were analyzed in a large group of VLBW and full-term Finnish children with the Finnish version of the CDI. There was no significant difference between the two groups in the size of the vocabulary at 2;0. Moreover, a clear gender difference was found in the lexicon size in full-term, but not in VLBW children. The composition analysis showed that the overall shapes of the trajectories for the main lexical categories as a function of vocabulary size were highly similar in the lexicons of the VLBW and full-term children, but in the vocabularies of >425 words there were differences in the percentages of nouns and grammatical function words between the two groups. The trajectories found in the lexicon of Finnish children were closely related to those described in the literature.

We found no significant difference between the VLBW and full-term children in their vocabulary size. This finding suggests that VLBW children as a group acquire their first lexicon quantitatively in a similar way to full-term children. The result of the recent longitudinal follow-up study by Jansson-Verkasalo (2003) supports this view. In this study no significant difference was found in the vocabulary size between the 17 VLBW and

17 matched controls at 2;0. However, it is also possible that the age correction used in the present study for VLBW children at 2;0 overestimates their true lexical capacity, as the vocabulary size of VLBW boys was larger than that of full-term boys. This difference was not statistically significant.

The other finding related to the lexicon size was that there was no gender difference in the size of the vocabulary in VLBW children, in contrast to the results for full-term children. This result is interesting when considering it in light of the findings in the literature. For example, Bornstein *et al.* (2004) reported a consistent advantage for healthy full-term girls over boys in expressive vocabulary development at 1;8, regardless of birth order or methodology. The fact that this trend was not found in VLBW children in the present study suggests that the early lexical development of prematurely born VLBW girls may be more affected than boys. Similar findings related to language skills in prematurely born children have been reported. Largo, Molinari, Cominale Pinto, Weber & Duc (1986) found that preterm boys born at 27–36 weeks of gestation performed slightly better than girls in sentence completion and grammar tests at 5;0 (see also Jennische & Sedin, 1999). However, for comparison, in a study by Sansavini *et al.* (2006), prematurely born males had a significantly smaller lexicon size than girls at 2;6. There was also a significant interaction between birth weight and gender: in the group of birth weight of $<$1000 g, boys produced significantly fewer words than girls (Sansavini *et al.*, 2006). VLBW boys and girls were not grouped according to birth weight in the present study, but there was a significant positive correlation between birth weight and vocabulary size in the group of all VLBW children. It is possible that boys and girls were distributed differently according to birth weight in the present study than in the study of Sansavini *et al.* (2006), and that this has caused a difference in the findings.

Maternal education was associated with the vocabulary size in VLBW, but not in full-term children in the present study. This result is in line with earlier findings. Maternal education has been reported quite consistently to have a positive effect on language (e.g. Menyuk *et al.*, 1995) and cognitive outcome (e.g. Brooks-Gunn, Klebanov, Liaw & Spiker, 1993) of prematurely born children. On the other hand, this effect is not always found in full-term children (Pan *et al.*, 2004). It may be that the more educated mothers are more sensitive to the needs of their high-risk children, and are thus more capable of supporting the development of language skills in the most adequate ways. Landry, Smith, Miller-Loncar & Swank (1997) reported that mothers who were sensitive to children's focus of interest and did not highly control or restrict their behavior, had children who had greater increases and faster rates of cognitive–language and social development. The relations were stronger in the group of high-risk

VLBW than in the group of low-risk VLBW or in full-term children. It was also found that higher levels of SES were related to a greater rate of increase in cognitive–language ages (Landry *et al.*, 1997).

The composition analysis revealed that VLBW children acquired different lexical categories in a roughly similar order and rate as did full-term children. The common nouns category was acquired most quickly, predicates followed that and the slowest growth was found in the grammatical function words category in both groups. Thus, the developmental sequence from routines→reference→predication→grammar in the lexicon (Caselli *et al.*, 1995) happened generally in a similar manner in VLBW and in full-term children. However, there were differences between the groups. Differences, although not statistically significant, were found at the first 50 words, where the percentage of social terms was higher and the percentage of common nouns lower in the lexicons of VLBW than in full-term children. In addition, there were no grammatical function words in the vocabularies of VLBW children with small lexicons, when the full-term children had some. Two significant differences were found in large vocabularies (>425): a higher percentage of common nouns and a lower percentage of grammatical function words in the lexicons of the VLBW than in the full-term children. These findings, especially when considering them from the point of view of the four-stage model of lexical development proposed in Bates *et al.* (1994) and Caselli *et al.* (1995, 1999), suggest that the differences found in the lexicon of the VLBW children might be developmental ones. It may be that VLBW children acquire different lexical categories at a slower rate than full-term children. However, if taking into consideration only the statistically significant differences found when the percentages of the different lexical categories were counted in two different ways, the picture is different. Then, the only difference between the two groups of children was in the percentage of grammatical function words in the lexicons of >425 words. This finding suggests that the language difficulties VLBW children have may not be at the level of the lexicon, but in grammar. The results of Jansson-Verkasalo (2003) support this hypothesis. In that study, it was found that VLBW children scored significantly lower on the language comprehension subtest of the Reynell Developmental Language Scales and had a significantly shorter maximum sentence length than control children at 2;0. Both of these values can be seen as measures of morphosyntactic skills. Furthermore, the finding of the present study related to the difference in the percentages of grammatical function words between the VLBW and full-term children is in line with the results of Koster *et al.* (2005). They found that children at risk for dyslexia had fewer closed class words and verbs in their vocabularies than controls at 1;7, when the vocabulary size was controlled. Our results also suggest that the acquisition of grammatical

markers in particular may be difficult for children at risk for language problems.

The composition analysis was done with the help of two different measures: we used the percentages counted from the total number of vocabulary items and the percentages counted from the opportunities afforded by the checklist within each category (Bates *et al.*, 1994; Caselli *et al.*, 1995). Both measures gave the same result on a general level: the order of appearance of different language categories was the same irrespective of which measure was used. The result is compatible with the results of Bates *et al.* (1994), who also used the same measures and reported a similar finding. Moreover, both measures found the same significant difference in the grammatical function words in large vocabularies, indicating that this difference between the two groups at 2;0 is clear and visible irrespective of the measure used. The other significant difference, the finding of the VLBW children having more common nouns in their large lexicons than full-term children, was found only when the percentages were counted from the size of the vocabulary. The trend was similar when the word opportunity percentages were used, but the difference was not significant. This result is somewhat more difficult to explain. It is not yet known what kind of predictive value the composition of the early lexicon has. More research is needed in order to clarify the meaning of this finding.

The overall shapes of the trajectories for different lexical categories in our Finnish-speaking samples are closely reminiscent of those described in the literature (Caselli *et al.*, 1999; Jackson-Maldonado *et al.*, 1993; Maital *et al.*, 2000). The result supports the universality of the four-stage model of lexical development set by Bates *et al.* (1994) and Caselli *et al.* (1999). The gradually increasing curve for grammatical function words in the lexicon of Finnish children, which is similar to Italian children but slightly different from English children (Caselli *et al.*, 1999), may suggest that the growth of different language units partly relates to the morphological structure of the target language. Grammar was not in the focus of the present study, but as Finnish is morphologically a rich language, as is Italian (Caselli *et al.*, 1999; Bornstein *et al.*, 2004), one may hypothesize that the linear growth of grammatical function words in the lexicon of Finnish children may reflect the greater morphological load of the target language. Because of this, children are bound to pay attention to these words more actively even at very early stages of vocabulary development, unlike for example, children growing up in an English-speaking environment (Caselli *et al.*, 1999).

The sample in the present study consisted of children who were all of the same age (compare Thordardottir, Weismer & Evans, 2002).Within this age group there was a high variation in the size and composition of the lexicon. Thus, the children in the sample with small lexicons were necessarily slow learners, and those with large lexicons were necessarily fast learners. It

should be stressed that the results of the present study are representative only for children at 2;0, and are not necessarily representative of all children, for example, those in the earliest stages of language acquisition.

The children in the present study were divided into two groups according to their cognitive level, and the lexicon size and composition of those VLBW and full-term children developing cognitively according to their age and those who were not, were compared. The group of VLBW children is a very heterogeneous one, and the cognitive impairments in this group are more common than in full-term children (e.g. Wolke & Meyer, 1999). The difference between the groups of VLBW and full-term children was also found in the present study: the minimum MDI value was 50 in the group of all VLBW children, while in the group of all full-term children it was 72. The group division according to the general cognitive level has also been used in other studies concentrating on language skills in prematurely born children. For example, Luoma *et al.* (1998) analyzed the speech and language skills of those prematurely born (born at $\leqslant$32 weeks of gestation) children without major neurological disability (low intelligence quotient included in criteria) separately in the cohort of five-year-old children (compare also Wolke & Meyer, 1999). For those working in clinical practice it is important to know how the language, or lexical, development is proceeding in VLBW children with normal cognitive development. Thus, the use of MDI subgroups made it possible to get more exact and clinically relevant information on the lexical development of VLBW children than if the group division had not been used. Furthermore, it should also be emphasized that, although the MDI has a verbal component, it cannot be considered a language test. A more specific evaluation is needed to get detailed information on the children's language skills.

Our results support the reliability of the structured maternal rating method. A high significant correlation was found between maternal ratings, the values of the CDI and the structured clinical measure, the MDI value on the BSID. Furthermore, the values for the full-term children in the present study are highly similar to those reported in the normative study of the Finnish CDI (*Mdn* 269, *M* 277·9, *S.D.* 162·7, min. 0, max. 595, Lyytinen, 1999) at 2;0. This gives even more support to the reliability and validity of the Finnish version of the CDI.

This study provides new information on the lexicon of VLBW and full-term Finnish children. In order to clarify the long-term relevance of the present findings, especially those related to VLBW children, there is a need for a longitudinal follow-up study where the language skills of VLBW children are assessed in detail. Based on our results, one can already note that because of the heterogeneity of the VLBW group, there is a need to pay attention to the development of their language skills in a more sophisticated way than to those of full-term children.

REFERENCES

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, S., Reilly, J. & Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language* **21**, 85–123.

Bayley, N. (1993). *The Bayley Scales of Infant Development*, 2nd ed. San Antonio, TX: Psychological Corporation.

Bornstein, M. & Cote, L. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean and American English. *Child Development* **75**, 1115–39.

Bornstein, M., Leach, D. & Haynes, O. (2004). Vocabulary competence in first- and second-born siblings of the same chronological age. *Journal of Child Language* **31**, 855–73.

Brooks-Gunn, J., Klebanov, P., Liaw, F. & Spiker, D. (1993). Enhancing the development of low-birthweight, premature infants: Changes in cognition and behaviour over the first three years. *Child Development* **64**, 736–53.

Caselli, M., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L. & Weir, J. (1995). A cross-linguistic study of early lexical development. *Cognitive Development* **10**, 159–99.

Caselli, C., Casadio, P. & Bates, E. (1999). A comparison of the transition from first words to grammar in English and Italian. *Journal of Child Language* **26**, 69–111.

Casiro, O., Moddemann, D., Stanwick, R., Panikkar-Thiessen, V., Cowan, H. & Cheang, M. (1990). Language development of very low birth weight infants and fullterm controls at 12 months of age. *Early Human Development* **24**, 65–77.

Dollaghan, C., Campbell, T., Paradise, J., Feldman, H., Janosky, J., Pitcairn, D. & Kurs-Lasky, M. (1999). Maternal education and measures of early speech and language. *Journal of Speech, Language and Hearing Research* **42**, 1432–43.

Fenson, L., Dale, P., Reznick, J., Bates, E., Thal, D. & Pethick, S. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development* **59**, 5 (Serial No. 242).

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj II (ed.), *Language development*, Volume 2: *Language, thought and culture*, 301–334. Hillsdale, NJ: Erlbaum.

Helasvuo, M.-L. (to appear). Aspects of the structure of Finnish. In A. Klippi & K. Launonen (eds), *Communication disorders in Finnish and in Finland*. Clevedon: Multilingual Matters.

Jackson-Maldonado, D., Thal, D., Marchman, V., Bates, E. & Gutierrez-Clellen, V. (1993). Early lexical development in Spanish-speaking infants and toddlers. *Journal of Child Language* **20**, 523–49.

Jansson-Verkasalo, E. (2003). Auditory event-related potentials as indices of language impairment in children born preterm and with Asperger syndrome. Unpublished doctoral dissertation, University of Oulu.

Jennische, M. & Sedin, G. (1999). Speech and language skills in children who required neonatal intensive care. II. Linguistic skills at 6, 5 years of age. *Acta Pediatrica* **88**, 371–83.

Kauschke, C. & Hofmeister, C. (2002). Early lexical development in German: a study on vocabulary growth and vocabulary composition during the second and third year of life. *Journal of Child Language* **29**, 735–57.

Klebanov, P., Brooks-Gunn, J. & McCormick, M. (1994). Classroom behavior of very low birth weight elementary school children. *Pediatrics* **94**, 698–708.

Korkman, M., Kirk, U. & Kemp, S. (1997). *A neuropsychological assessment of children* [in Finnish], 1st ed. Helsinki, Finland: Psykologien kustannus.

Koster, C., Been, P., Krikhaar, E., Zwarts, F., Diepstra, H. & Van Leeuwen, T. (2005). Differences at 17 months: Productive language patterns in infants at familial risk for dyslexia and typically developing infants. *Journal of Speech, Language and Hearing Research* **48**, 426–38.

Landry, S., Smith, K., Miller-Loncar, C. & Swank, P. (1997). Predicting cognitive-language and social growth curves from early maternal behaviors in children at varying degrees of biological risk. *Developmental Psychology* **33**, 1040–53.
Landry, S., Smith, K. & Swank, P. (2002). Environmental effects on language development in normal and high-risk child populations. *Seminars in Pediatric Neurology* **9**, 192–200.
Largo, R., Molinari, L., Cominale Pinto, L., Weber, M. & Duc, G. (1986). Language development of term and preterm children during the first five years of life. *Developmental Medicine and Child Neurology* **28**, 333–50.
Luoma, L., Herrgård, E., Martikainen, A. & Ahonen, T. (1998). Speech and language development of children born at <32 weeks'gestation: A 5-year prospective follow-up study. *Developmental Medicine & Child Neurology* **40**, 380–87.
Lyytinen, H., Ahonen, T., Eklund, K., Guttorm, T., Laakso, M.-L., Leinonen, S., Leppänen, P., Lyytinen, P., Poikkeus, A.-M., Puolakanaho, A., Richardson, U. & Viholainen, H. (2001). Developmental pathways of children with and without familial risk for dyslexia during the first years of life. *Developmental Neuropsychology* **20**, 535–54.
Lyytinen, P. (1999). Varhaisen kommunikaation ja kielen kehityksen arviointimenetelmä. Jyväskylän yliopiston Lapsitutkimuskeskus ja Niilo Mäki instituutti. Jyväskylä: Niilo Mäki Säätiö.
Lyytinen, P. & Lyytinen, H. (2004). Growth and predictive relations of vocabulary and inflectional morphology in children with and without familial risk for dyslexia. *Applied Psycholinguistics* **25**, 397–411.
Nieminen, P. (1991). Äidin ja lapsen kommunikaatio ja lapsen kielen omaksuminen. Unpublished doctoral dissertation, University of Tampere (ser A, vol. 323).
Maital, S., Dromi, E., Sagi, A. & Bornstein, M. (2000). The Hebrew Communicative Development Inventory: language specific properties and cross-linguistic generalizations. *Journal of Child Language* **27**, 43–67.
Menyuk, P., Liebergott, J. & Schultz, M. (1995). *Early anguage development in full-term and in premature infants*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
Mikkola, K., Ritari, N., Tommiska, V., Salokorpi, T., Lehtonen, L., Tammela, O., Pääkkönen, L., Olsen, P., Korkman, M. & Fellman, V. (2005). Neurodevelopmental outcome at 5 years of age of a national cohort of extremely low birth weight infants who were born in 1996–1997. *Pediatrics* **116**, 1391–400.
Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development* **38** (Serial No. 149).
Oliver, B., Dale, P. & Plomin, R. (2004). Verbal and nonverbal predictors of early language problems: An analysis of twins in early childhood back to infancy. *Journal of Child Language* **31**, 609–631.
Pan, B., Rowe, M., Spier, E. & Tamis-LeMonda, C. (2004). Measuring productive vocabulary of toddlers in low-income families: Concurrent and predictive validity of three sources of data. *Journal of Child Language* **31**, 587–608.
Paul, R. (2001). Predicting outcomes of early expressive language delay: Ethical implication. In D. Bishop & L. Leonard (eds), *Speech and language impairments in children. Causes, characteristics, intervention and outcome*, 2nd ed, 195–209. East Sussex: Psychology Press Ltd.
Pine, J., Lieven, E. & Rowland, C. (1996). Observational and checklist measures of vocabulary composition: What do they mean? *Journal of Child Language* **23**, 573–89.
Rescorla, L. (2005). Age 13 language and reading outcomes in late-talking toddlers. *Journal of Speech, Language and Hearing Research* **48**, 459–72.
Riitesuo, A., (2000). A Preterm Child Grows. Unpublished doctoral dissertation, University of Jyväskylä.
Rvachew, S., Creighton, D., Feldman, N. & Sauve, R. (2005). Vocal development of infants with very low birth weight. *Clinical Linguistics & Phonetics* **19**, 275–94.
Sansavini, A., Guarini, A., Alessandroni, R., Faldella, G., Giovanelli, G. & Salvioli, G. (2006). Early relations between lexical and grammatical development in very immature Italian preterms. *Journal of Child Language* **33**, 199–216.

Thordardottir, E., Weismer, S. & Evans, J. (2002). Continuity in lexical and morphological development in Icelandic and English-speaking 2-year-olds. *First Language* **22**, 3–28.
Toivainen, J. (1997). The acquisition of Finnish. In Dan Slobin (ed.), *The crosslinguistic study of language acquisition*, Volume 4, 87–182. Mahwah, NJ: Lawrence Erlbaum.
Weismer, S. (2001). Intervention for children with developmental language delay. In D. Bishop & L. Leonard (eds), *Speech and language impairments in children. Causes, characteristics, intervention and outcome*, 2nd ed, 57–176. East Sussex: Psychology Press Ltd.
Wolke, D. & Meyer, R. (1999). Cognitive status, language attainment, and prereading skills of 6-year-old very preterm children and their peers: The Bavarian Longitudinal Study. *Developmental Medicine & Child Neurology* **41**, 94–109.